
MASSACHUSETTS
INSTITUTE of TECHNOLOGY
LIBRARIES

working paper department of economics

FORMAL AND REAL AUTHORITY IN ORGANIZATIONS

Philippe Aghion
Jean Tirole

95-8 **June, 1994**

massachusetts institute of technology

50 memorial drive
cambridge, mass. 02139

FORMAL AND REAL AUTHORITY IN ORGANIZATIONS

Philippe Aghion
Jean Tirole

95-8 **June, 1994**

Formal and Real Authority in Organizations*

Philippe Aghion† and Jean Tirole‡

June 1, 1994

*The authors are grateful to Oliver Hart and Martin Hellwig for helpful discussions, to Patrick Bolton, Leonardo Felli, and especially Denis Gromb and David Martimort for helpful comments on a first draft, and to the Centre National d'Etudes des Télécommunications for financial support.

†Nuffield College, Oxford and EBRD, London

‡IDEI, Toulouse, France; CERAS, Paris; and MIT, Cambridge, MA USA

Abstract

The paper develops a theory of the separation between formal authority (the right to decide) and real authority (the effective control over decisions), and illustrates how a formally integrated structure can accommodate various degrees of "real" integration. Real authority is determined by the structure of information, which in turn depends on the allocation of formal authority. An increase in an agent's real authority promotes initiative but results in a loss of control for the principal.

After showing that firm boundaries and information structures are intertwined, the paper examines a number of factors that increase the subordinates' real authority in an integrated structure: overload, lenient rules, urgency of decision, reputation, performance measurement, and multiplicity of superiors.

Lastly, the amount of communication in an organization is shown to depend on the allocation of formal authority.

1 Introduction

Over forty years ago, Herbert Simon defined authority as the right to select actions affecting part or the whole of an organization.[1] As pointed out by Grossman-Hart (1986) and Hart-Moore (1990), authority may be conferred by the ownership of an asset, which gives the owner the right to take decisions concerning the use of this asset. Authority may more generally result from an explicit or implicit contract allocating the right to decide on specified matters to a member or group of members of the organization.[2]

This *formal authority* however need not confer *real authority*, that is an effective control over decisions, upon its holder. For example, it is commonplace to observe that shareholders have limited control over their board of directors, which itself may be subject to the domination of the top executives, who in turn often rubber-stamp the divisions' projects, and so forth. Similarly, the president of a country really controls only a small number of the decisions taken by the executive branch. This paper develops a theory of the separation between formal authority and real authority,[3] and illustrates how a formally integrated structure can accommodate various degrees of "real" integration.

Our approach follows Max Weber's (1968) description of "rational" or "legal" authority. Weber notes that officials, employees, and workers attached to the administrative staff of a bureaucracy do not themselves own the non-human means of production and administration, yet they may exert substantial control over the bureaucratic machinery (p217-225). For example, he observes that :

> "Bureaucratic administration means fundamentally domination through knowledge. This is the feature of it which makes it specifically rational. This consists on the one hand in technical knowledge which, by itself, is sufficient to ensure it a position of extraordinary power. But in addition to this, bureaucratic organizations, or the holders of power who make use of them, have the tendency to increase their power still further by the knowledge growing out of experience in the service. For they acquire through the conduct of office a special knowledge of facts and have available a store of documentary material peculiar to themselves. While not peculiar to bureaucratic organizations, the concept of "official secrets" is certainly typical of them. It stands in relation to technical knowledge in somewhat the same position as commercial secrets do to technological training. It is

[1] "We will say that the boss exercises authority over the worker if the worker permits the boss to select x [a "behavior", i.e., any element of the set of specific actions that the worker performs on the job]. That is, the worker accepts authority when his behavior is determined by the boss's decision. In general, the worker will accept authority only if x_0, the x chosen by the boss, is restricted to some subset (the worker's "area of acceptance") of all the possible values. This is the definition of authority that is most generally employed in modern administrative theory" [Simon (1951,p294)].

[2] This corresponds to Weber's (1968) notion of "rational" or "legal" authority. Weber distinguishes among three types of authority, the other two being "traditional authority" and "charismatic authority".

[3] The distinction between ownership and control is reflected in the principal-agent literature, starting with Mirrlees (1975), Holmström (1979) and Shavell (1979). This literature studies the structuring of compensation and rewards, assuming full control by the agent, and does not investigate the allocation of authority, formal or real.

a product of the striving for power."

As in Weber, the key to our analysis of formal vs real authority is asymmetric information. The superior can always reverse her subordinate's decision, but will refrain from doing so if the subordinate is much better informed and if their objectives are not too antinomic. We formalize this idea in a straightforward way. The subordinate exerts effort (shows initiative) to suggest a project to the superior. The latter as well chooses how much to learn about the potential project. Once informed, the subordinate recommends a project that sometimes is not optimal for the superior, because from the point of view of the agent this project creates a higher private benefit, yields better career opportunities, or requires less effort to be implemented than the optimal project. Formal authority prevails when the superior is informed, since the superior then chooses her preferred project (which may or may not coincide with the subordinate's proposal). In contrast, a poorly informed superior optimally rubber-stamps the subordinate's proposal by fear of picking a worse alternative. The subordinate then has real, although no formal authority.

A superior who for instance is overloaded and therefore has little time to monitor her subordinates' decisions loses control and involuntarily endorses many suboptimal projects. Conversely, too much information also hurts the superior. The very prospect of influencing decisions is what creates *initiative* from the subordinate. The subordinate loses motivation when the superior's incentives (for instance, idleness, obsession or high competency) lead him to "stand constantly behind the subordinate's shoulders", that is to monitor very thoroughly. There is thus a general trade-off between initiative and loss of control. Furthermore, the degree of separation of ownership and control is endogenous; it depends on the nature of the task, on who performs and monitors it, and on the organizational structure.[4]

This new approach to ownership and control based on the distinction between formal and real authority delivers a number of interesting conclusions.

First, it provides a logical relationship between the information structure and the allocation of decision rights within an organization:[5] The reallocation of formal authority to the agent prevents the principal from overruling the agent in those situations (à la Grossman-Hart (1986)) where both parties have acquired the relevant information about the potential projects' payoffs. The transfer of formal authority to the agent thus credibly increases the agent's initiative. On the other hand, the principal has more incentives to become informed

[4]A straightforward application of this idea is the prediction that the delegation of day-to-day activities is more likely to be observed than the delegation of key investments. This is not due to a relative irrelevance of day-to-day activities, but rather to the fact that they are very numerous and therefore more costly to monitor.

[5]Note also that an agent who is "outside the organization" in the sense of not having signed an initial contract with the principal, may also exert "real authority", in that he may bring information that influences decision making.

if she has the right to reverse decisions. *Firm boundaries and information structures are thus intertwined*, and the size of a firm is ultimately determined by the above trade-off between the subordinate's initiative and the principal's cost of losing control over the choice of projects.

Second, whilst conferring decision rights on the agent increases his initiative, there are also a number of *factors that encourage initiative when the principal retains formal authority.* One factor of initiative and loss of control is a wide span of control, which raises the principal's marginal cost of monitoring each agent. We show that there is a sense in which optimal organizations always function in a situation of overload. Another raise-the-monitoring-cost way of committing not to stifle initiative is to refrain from imposing rules that constrain the agent within a set of easily monitorable activities. Alternatively, the gain from intervention can be reduced by spreading its benefits among several principals/owners;[6] intervention can also be made more difficult by splitting property rights among several superiors (as in the case of a matrix organization or of multiministry oversight) and by requiring that intervention be unanimously agreed upon.

Other factors that encourage initiative include the urgency of decision making, which does not give the superior much time for a thorough investigation; repeated interaction, which allows the superior to develop a reputation for not intervening in matters that are fairly inconsequential to her and for intervening only in important matters; and improved performance measurement.

Third, our approach enables us to provide a modest, but first, step toward the integration of "collective bounded rationality" and incentives. A recent literature has studied the implications of imperfect communication on the functioning of an organization.[7] This research has taken the imperfection of communication as given and in general has assumed that members of the organization pursue the same objective. We would like to go one step beyond by *endogenizing the limits to communication.* The basic idea is to depart from the traditional team theoretic framework of the literature to allow members to have dissonant objectives. The communication of information is then strategic and depends on the authority relationship. In particular, less communication may take place if the principal has formal authority because the agent is concerned that the principal might abuse her authority once she is well informed. This will typically be the case if the principal's and subordinate's objectives are sufficiently dissonant. In the opposite case where these objectives are sufficiently congruent we show that communication may instead be encouraged by the agent's subordination to the principal.

[6]This theme is developed in a corporate finance framework by Burkart, Gromb and Panunzi (1994), with a number of interesting implications.

[7]See, e.g., Marschak-Radner (1992), Sah-Stiglitz (1986), Bolton-Dewatripont (1992) and Radner (1992).

Although our analysis is, to the best of our knowledge, new, it makes use of building blocks developed by other authors. The moral hazard and property rights literatures supplied the two polar cases. A seminar given by Diego Rodriguez and Dimitri Vayanos at MIT in 1991 contained several seeds of the basic model described here.[8] Papers by Crémer (1992), Riordan (1990), and Schmidt (1991) develop in different contexts the idea that too much information may hurt a principal.[9] Riordan (1990) argues that information allows principals to expropriate the agents' specific investments. He provides a definition of vertical integration based on information. Crémer (in the context of a corporation) and Schmidt (in a paper on privatization) show that poor information allows principals to avoid (ex ante costly) renegotiation of long term contracts with agents. In Crémer's paper, the principal publicly chooses the accuracy of a technology used to monitor the agent's type. A more accurate technology reduces the agent's incentive to work to signal high ability. Riordan and Schmidt discuss property rights and, in the tradition of Arrow (1975), posit (but do not formally establish) a link between property rights and information structure.

The paper is organized as follows: Section 2 presents the model. Section 3 analyzes it: It first relates the information structure and equilibrium payoffs to the allocation of formal authority; it then demonstrates a tight link between our incomplete contract/authority methodology and a complete contract approach. Section 4 looks at factors favoring the agent's initiative when the principal has formal authority (overload, lenient rules, urgency of decision, reputation, better performance measurement, multiple principals), and derives implications for business management. Section 5 endogenizes communication, and section 6 concludes.

2 The model

A hierarchy composed of a principal (she) and an agent (he) can implement one or zero project. The principal hires the agent to collect information about and implement the project. Examples of hierarchies we have in mind are board of directors or trustees/management, CEO/division manager, thesis advisor/student, foreman/worker, or supranational authority/country.

• *Projects*: The agent screens among $n \geq 3$ potential and a priori identical projects on behalf of the principal. To each project $k \in \{1, \cdots, n\}$ is associated a verifiable monetary gain or profit B_k for the principal and a private benefit b_k for the agent. [These payoffs are gross of any monetary transfer between the two parties]. The agent's private benefit includes perks on the job, acquisition of human capital, the possibility of signalling ability, or (minus)

[8] Their 1993 discussion paper focuses on different themes than the ones considered here.

[9] The literature on the ratchet effect also emphasizes a cost for a principal from being well informed.

the disutility of implementing the project.[10] If no project is implemented, the profit and the private benefit are both equal to zero. "No project" can formally be treated as project 0, with known payoffs $B_0 = b_0 = 0$.

Only two of the n potential projects are actually "serious" or "relevant". All other projects yield "sufficiently negative" payoffs to both parties. Anticipating somewhat, this will imply that an uninformed agent prefers to confess ignorance and to recommend inaction rather than to recommend a specific project, and that similarly an uninformed principal would not choose by herself to undertake a project.

One of the two relevant projects yields profit $B > 0$ to the principal and the other 0. Similarly, one of the two relevant projects yields private benefit $b > 0$ to the agent, and the other 0. The ex ante probability that the same project is preferred by both is $\alpha \in [0, 1]$, the parameter of congruence.[11]

• *Preferences*: The principal is risk neutral and has utility $B_k - w$ if project k is chosen, and w is the wage paid to the agent. The agent is protected by limited liability, so $w \geq 0$. The agent's utility is then $u(w) + b_k$, where $u(\cdot)$ is increasing and concave.

For expositional simplicity we will assume that the agent is infinitely risk averse to income. He therefore does not respond to monetary incentives and receives a constant wage equal to his reservation wage of zero. Alternatively, the agent may not be infinitely risk averse, but the principal's benefit is noncontractible; the agent's wage is then again a constant. [Section 4.5 shows that the model can be straightforwardly extended to allow the agent to respond to monetary incentives. Profit sharing then lowers the principal's and raises the agent's utility from picking a profitable project.]

• *Information*: The agent acquires information in a binary form. At private cost $g_A(e)$, he perfectly learns the payoffs of all candidate projects with probability e. With probability $(1 - e)$, the agent learns nothing and still views the projects as identical.

[10] For example, a common complaint in large conglomerates is that R&D operations get out of control and R&D units put too much emphasis on basic research ("R") and too little on developing commercial applications ("D"). The emphasis on basic research stems from the units' private benefit.

[11] This congruence parameter will be treated as *exogenous* in the following analysis. However, one could think of various methods whereby the principal might affect congruence with her subordinates: for example, investments in the recruiting and training of new employees, design of career profiles, enforcement of [contractual] rules restricting the subordinates' set of possible actions.

An interesting determinant of α is the possibility that other agents working for the principal are affected by the activity of this agent. A case in point is supplied by the difference in behavior of IBM and Fujitsu vis-à-vis their units. IBM is less successful in keeping an arms-length relationship with its units than Fujitsu with its British subsidiary ICL. According to *The Economist* (Management Focus. April 10, 1993), one reason is that Fujitsu and ICL are not serious rivals anywhere while IBM's units are often competitors. Under the plausible hypothesis that the existence of externalities on other units reduces the congruence parameter, this view is consistent with our theory, which predicts that the subordinate has more discretion, the higher the congruence.

A systematic exploration of the various ways of endogenizing the congruence parameter lies beyond the scope of this paper.

Similarly, the principal chooses how much time or effort to devote to learning payoffs. At private cost $g_P(E)$, she becomes perfectly informed about the payoffs with probability E and learns nothing with probability $(1-E)$.

The principal's acquisition of information can be contemporaneous with the agent's or else start after the agent makes his report. We will refer to these two possibilities as the *simultaneous* and *sequential* models, respectively. Which variant is more relevant depends on the context. Sequential investigations usually are less time consuming for the principal, who can already build on an existing report. On the other hand, the principal may not want to wait until the report accrues to start her investigation, as otherwise she may be forced to accept the agent's proposal by lack of time.[12] Because the simultaneous and sequential cases yield essentially the same results, we will focus on the simultaneous case and content ourselves with illustrating the sequential case in section 4.3. We leave the endogenization of the timing (simultaneous vs sequential) for future research.

The disutilities of effort $g_A(\cdot)$ and $g_P(\cdot)$ are increasing and strictly convex and satisfy $g_i(0)=0$, $g_i'(0)=0$, $g_i'(1)=\infty$, $i=A,P$.

• *Communication*: In most of the paper we can assume that information is either *hard* or *soft*. Hard information about a project's payoffs can be costlessly and instantaneously verified by the other party if communicated by the party who collected it. Soft information cannot be verified by the other party and therefore its communication must be interpreted as a pure suggestion for a project choice. The specific results of section 4.3 rely on the information being soft.

• *Authority*: In the case of *P-formal authority* (which we will occasionally label "integration" and is the main focus of this paper), the principal has the *formal* authority and is called the "*superior*". The principal may always overrule the agent (the "*subordinate*"). She indeed does so if she is informed and if the agent's recommendation is not "congruent". In this case, the principal has both the formal and real authority over the choice of project and can fully dispense with the agent's information and recommendation. Otherwise, she (optimally) rubber-stamps the agent's proposal. We will then say that the agent has *real* authority.

Our payoff structure implies that there is no need to include an "exit option" for the subordinate; for, the superior always takes a decision that yields nonnegative utility to both. The standard institution of letting subordinates quit if they are unhappy with their superiors' decisions emerges naturally in the variant of our model in which the principal's preferred

[12]For instance, directors of a company or external members of a thesis jury are usually forced to rubber-stamp the annual report or to accept the thesis if they have waited until receiving the documents to become involved.

project may impose a substantial loss of utility for the agent.

Under *A-formal authority* (which we will occasionally label "delegation"), the "*independent agent*", if informed, picks his preferred project and cannot be overruled by the principal. That is, the agent now has formal authority. Note that this covers the situation in which the agent is an employee who contractually receives an irrevocable right to take this particular decision.[13]

We will assume that whoever has authority prefers choosing a project that gives him zero, but benefits the other party to doing nothing (equivalently, we could assume that the two relevant projects yield strictly positive payoffs to both parties).

• *Contracts.* In most of the analysis, we adopt the incomplete contracting approach (Grossman-Hart (1986)) by positing that projects cannot be described and contracted upon ex ante. The initial contract only specifies an allocation of *formal authority* (control rights) to one of the two parties.

The *timing* is as follows: (i) The principal proposes a contract that allocates *formal* authority to her or the agent over the future choice of projects;[14] (ii) the parties privately gather information about the n projects' payoffs; (iii) the party who does *not* have formal authority communicates a subset of the relevant projects' payoffs he has learned to the controlling party; given that there are only two relevant projects, this subset may comprise zero, one or two project payoffs descriptions; (iv) the controlling party picks a project (or none) on the basis of his information and the information communicated by the other party.

To put the incomplete contracting approach somewhat in perspective, section 3.4 explores the polar assumption that projects can be described and contracted upon ex ante, even though their payoffs are ex ante unknown to both parties. We will see that under specific circumstances, the optimal complete contract corresponds exactly to a (possibly random) authority allocation scheme, and we will point at the implications of contract incompleteness when these circumstances do not obtain.

• *Payoffs under the two allocations of authority.* Under *P-formal authority* (integration),

[13]The allocation of authority must therefore be thought of as issue specific. For instance, the chairperson of a department can pick the salary or the teaching load of a professor but has no right to choose the professor's research agenda. Another case in point is that of Du Pont in 1921, in which "headquarters did not have the [formal] authority to step in and interfere with divisional decisions unless something was clearly going wrong" (Chandler, in Continental Bank (1993), p56, bracket added). Our model has a single project/issue, but we find its extension to multiple decision rights and clusters of authority potentially fascinating.

Note also that the delegation of formal authority is sometimes partial in that the principal may be able to regain authority at a large cost: An official may be impeached, or a board of directors overruled by shareholders through a takeover or a proxy fight.

[14]That is, we assume that there is ex ante a competitive supply of potential agents, so that the allocation of authority between the two parties is the one that maximizes the principal's ex ante expected utility.

the utilities are

$$u_P = EB + (1 - E)e\alpha B - g_P(E) \tag{1}$$

and

$$u_A = E\alpha b + (1 - E)eb - g_A(e). \tag{2}$$

That is, with probability E, the principal is informed and picks her preferred project. With probability $(1 - E)$, the principal is uninformed. With probability e, the agent is informed and suggests his preferred project. The principal then either learns from his recommendation the payoffs attached to this project (hard information) or is still uncertain about whether the agent proposes her preferred project (soft information). Either way, the principal optimally rubber-stamps the agent's proposal.[15]

Under *A-formal authority* (delegation) when informed, the agent simply chooses his preferred project. When the agent is uninformed, and the principal is informed, the principal suggests her preferred project, which is then implemented by the agent. So, using the superscript "d" for "delegation", preferences are

$$u_P^d = e\alpha B + (1 - e)EB - g_P(E) \tag{3}$$

and

$$u_A^d = eb + (1 - e)E\alpha b - g_A(e). \tag{4}$$

Note that the agent's lack of responsiveness to monetary incentives precludes any renegotiation of the exercise of authority.

Remark: That the principal's formal authority becomes entirely ineffective when the principal is uninformed follows from the specific projects payoff structure, in particular from the principal's weakly preferring a (relevant) noncongruent project to no project at all. To see this, consider the following example: there are three relevant projects (thus $n \geq 4$), say $k = 1, 2$ and 3. Project 3 yields a strictly negative profit to the principal, whilst projects 1 and 2 yield (as above) a nonnegative profit. Then, if information is hard, the principal whenever uninformed can still use her formal authority in order to elicit information about project 3 vs projects $\{1, 2\}$ from the agent. The principal will then rubber-stamp the agent's decision only conditionally on information ruling out the negative-profit project having been disclosed to her.

[15]Note that the superior is always better off, the higher the agent's effort. This need not hold in a multi-task extension of our model; for, the agent's effort on this task may then crowd out his effort on other tasks. Consequently, the agent's initiative in one task may not be positively valued by the principal.

3 Analysis

3.1 The reaction curves under P-formal authority

The first-order conditions when the principal has formal authority are:

$$(1 - e\alpha)B = g'_P(E) \tag{5}$$

and

$$(1 - E)b = g'_A(e). \tag{6}$$

The principal supervises more, the higher her stake and the lower the congruence parameter and the agent's effort. The agent *demonstrates more initiative*, the higher his private benefit and *the lower the principal's interference.*

We assume that the two systems of equations {(5),(6)} has a unique, stable intersection $(E.e)$[16] [Such an assumption is not needed in the sequential case: Because the principal acquires information only if the agent makes a proposal, E is independent of e and the stability condition is automatically satisfied.[17]]

The fact that the agent's reaction curve (6) is *downward* sloping is a crucial feature of this (or any) delegation model. In contrast, if the agent's reaction curve were *upward* sloping, the principal would never want to reduce her degree of interference E for strategic reasons. This latter case might correspond either to a situation of *strategic complementarity* if the principal's reaction curve were also upward sloping or to a *supervision* situation if the principal's reaction curve were downward sloping.

Finally, note the first-order condition (5) implies that the principal's reaction to the agent's effort is also downward sloping. The interpretation is that the principal can and has an incentive to substitute for the agent in case the agent does not work. Suppose in contrast that the agent is indispensable. That is, the principal cannot invent a project if the agent does not come up with a proper suggestion; on the other hand, the principal can carefully read any file gathered by the agent and modify the existing project. The principal's reaction curve is then likely to slope up. [The comparative statics results obtained in this paper carry over to this case for shifts in the principal's reaction function, but not for shifts in the agent's reaction function.]

[16] That is, $\alpha bB < g''_P(E)g''_A(e)$.

[17] For example, with soft information, the principal's payoff in the sequential case is:

$$u_P = e[EB + (1 - E)\alpha B - g_P(E)].$$

3.2 The basic tradeoff between loss of control and initiative

Suppose that for some "exogenous" reason (e.g., because of overload), the marginal cost of effort of the principal (i.e., g'_P) increases. The effect on the principal's expected payoff is a priori ambiguous. On the one hand, ceteris paribus, the principal's probability of becoming informed about the projects' payoffs (E) decreases (see equation (5)); the principal thus loses *real* authority (i.e., control) over the choice of project, with a higher resulting risk of having to endorse suboptimal projects. On the other hand, the reduction in the principal's intervention E encourages *initiative* from the subordinate (see equation (6)), which in turn raises the principal's expected (monetary) benefit.

This basic tradeoff between *loss of control* and *initiative* determines the optimal allocation of *real* authority when the initial contract allocates *formal* authority to the principal. The question then arises as to how the principal can actually *commit* herself to delegate an (optimal) amount of real authority to her subordinate while retaining formal authority. This question is taken up in detail in section 4. One (extreme) way for the principal to foster the agent's initiative is to relinquish her *formal* authority, as we shall now see.

3.3 The optimal allocation of *formal* authority

Whether *formal* authority should optimally be allocated to the principal or to the agent hinges on the tradeoff between loss of control and initiative. More formally, when the agent has formal authority, the first-order conditions are:

$$(1-e)B = g'_P(E) \tag{7}$$

and

$$(1-\alpha E)b = g'_A(e). \tag{8}$$

Assuming again that {(7), (8)} yields a unique, stable equilibrium (E^d, e^d),[18] one can show that :

$$E > E^d \quad \text{and} \quad e < e^d.$$

Our model thus explains why *the absence of integration (where "integration" is understood as P-formal authority) corresponds*, as is often suggested in the literature (e.g., in Williamson (1975,1985)), *to an arms-length relationship.* Because the principal cannot overrule the agent, she has less incentive to become informed and thus giving the formal authority to the agent is a credible way for the principal to commit not to intervene. *The cost of leaving initiative to the agent*, on the other hand, is *a loss of control.* The agent takes suboptimal decisions more often.

[18]The stability condition is the same as in the case of P-formal authority.

Example: Let us compare the two governance structures in the quadratic cost specification, in which $g_A(e) = e^2/2$ and $g_P(E) = E^2/2$.[19]

One has

$$E = \frac{B(1-\alpha b)}{1-\alpha bB}, \quad e = \frac{b(1-B)}{1-\alpha bB},$$

and

$$E^d = \frac{B(1-b)}{1-\alpha bB}, \quad e^d = \frac{b(1-\alpha B)}{1-\alpha bB} \quad .$$

Furthermore,

$$u_P - u_P^d = \frac{B^2 b(1-\alpha)}{(1-\alpha bB)^2}[1 - \frac{b}{2}(1+\alpha) - \alpha(1-\alpha bB)].$$

P-formal authority is optimal for low levels of congruence (α small), and A-formal authority dominates when preferences almost coincide (α close to 1). Indeed, there is a cutoff value $\alpha^* \in (0,1)$ such that P-formal authority is optimal if and only if $\alpha < \alpha^*$.[20]

3.4 Complete vs incomplete contracts

The purpose of this subsection is to step back and reflect on our methodological approach.[21] One of the common motivations for using incomplete contracting models is that the standard complete contract paradigms (moral hazard, adverse selection or Nash implementation) fail to account for notions such as ownership and authority. We have adopted an incomplete contracting approach to analyze authority and its delegation and found it very useful to conceptualize our intuitions and obtain others. Yet, in view of the current lack of proper foundations for the incomplete contracting methodology, an investigation of what could be achieved with complete contracts will bring us a deeper perspective.

Interestingly, while the notion of authority finds its natural habitat in an incomplete contracting framework, it does not rely on such an interpretation. As we have seen, choosing who should decide amounts to answering two subsidiary questions. First, who is informed? A party with private information can manipulate decision making by revealing a coarser information structure than his structure. More concretely, he will typically disclose his preferred project and abstain from revealing information that may lead to the adoption of projects he likes less. Second, who gets his way when both parties are informed? The answer to this second question hinges on the classic tradeoff between ex post efficiency of

[19]This case does not satisfy the assumption that $g_i'(1) = \infty$, but is still amenable to our analysis as long as we assume that, in the relevant range of parameters, the parties do not become informed with probability 1.

[20]To prove this, one uses the fact that the probabilities e and E must be less than one. One can also show that the agent prefers having formal authority when congruence is low, but may prefer not to have formal authority in specific circumstances (high congruence, private benefit small relative to the principal's payoff.)

[21]This section is not required reading for the rest of the paper, and may be skipped by readers who are primarily interested in the incomplete contracting approach.

decision making and ex ante incentives to acquire information. These two questions surface as well in a world of complete contracting with multiple moral hazard. We now state the circumstances under which the outcome can alternatively be interpreted as resulting from an optimal complete contract.

• In a first step, we maintain our assumption that the principal and the agent are the only two parties (or that B is noncontractible); as we shall see, this assumption essentially rules out the use of the sharing of the principal's profit with outsiders in order to encourage the agent's initiative. We assume limited liability for both parties.[22]

The timing is as follows: (i) The two parties sign a contract; (ii) they privately gather information about the n projects' payoffs (for simplicity, a party who has learned the payoffs also knows whether the other party has learned them as well); (iii) they send messages; (iv) they possibly renegotiate the initial contract; (v) the initial contract (if still in force) or the new contract (if renegotiation has taken place) is implemented. A *message* m_i, $i = A, P$, is either $\emptyset$ ("I have not learned the payoffs"), or a pair of payoffs for the principal and the agent for each project $k = 1, \cdots, n$ ("I have learned the payoffs, and they are..."). A *contract* specifies probabilities[23] $p_k(m_P, m_A)$ of implementing project k, such that

$$\sum_{k=1}^{n} p_k(m_P, m_A) \leq 1.$$

There are three states of the world in which at least one party learns the payoffs and therefore a project is to be implemented: 1: only the principal learns the payoffs (probability $E(1-e)$); 2: only the agent learns the payoffs (probability $(1-E)e$); 3: both learn the payoffs (probability Ee). In state $j \in \{1,2,3\}$, let x_j denote the equilibrium probability that the congruent project is implemented if there indeed is congruence; in case of noncongruence, the principal's preferred project is implemented with equilibrium probability y_j and the agent's preferred project is implemented with equilibrium probability z_j, with $y_j + z_j \leq 1$.

We assume that the renegotiation process is a finite bargaining game. In each subperiod $\tau = 1, \cdots, T$ of this process, one of the parties suggests a project or doing nothing (or more generally, a probability distribution over these decisions); the other party accepts or turns down the offer. In the latter case, bargaining proceeds to the next subperiod. The decision specified by the initial contract for the messages sent at stage (iii) is implemented if no agreement is reached by subperiod T. Furthermore an informed party can disclose payoffs about particular projects at the start of each subperiod. For concreteness we will assume

[22] Alternatively, the two parties might be infinitely risk averse under zero income. An arbitrarily small probability of error would then endogenize the limited liability assumption.

[23] Unlike in conventional mechanism design, there is no leeway in using monetary transfers to adjust incentives because there are only two parties, profits cannot be thrown away (since renegotiation is feasible), and one of the parties, the agent, does not respond to monetary incentives.

that information is hard. The argument can be extended with minor modifications to soft information.

The possibility of renegotiation[24] implies that $x_i = 1$ for all i. Furthermore, incentive compatibility plus the possibility of renegotiation imply that $y_1 = 1$ and $z_2 = 1$. That is, in case of asymmetric information and noncongruence, the informed party can guarantee himself his preferred payoff by disclosing only his preferred project, which, because of renegotiation, is then implemented with probability 1. To see this, note first that, because irrelevant projects yield very negative payoffs, when both announce they have not learned the payoffs, the initial contract yields status quo expected payoffs equal to zero (if with probability one no project is implemented) or negative (otherwise.) Second, suppose that the informed party feigns ignorance ($m_i = \emptyset$, $i = A, P$) at stage (iii). Now consider the stage (iv) bargaining game and suppose that the informed party has disclosed only his preferred project at the last subperiod T. Then, whoever makes the last offer necessarily proposes this project and, assuming that when indifferent a party chooses what is preferred by the other party (alternatively we could assume that each relevant project yields at least a small positive amount to each party), the other party accepts. Backward induction then shows that by disclosing only his preferred project, the informed party necessarily gets his own way. Last, the possibility of renegotiation implies that $y_3 + z_3 = 1$. Who gets his own way when the information is symmetric (state 3) is determined by the relative desirability of the investments e and E.[25]

To sum up, *the optimal contract is a (random) authority scheme*, in which authority is ex post conferred upon the principal with probability y_3 and upon the agent with probability z_3 (note that the allocation of authority was shown in section 3.1 to be irrelevant in states of asymmetric information, in which the informed party always gets his own way). Because random authority mechanisms are allowed by the incomplete contracting approach

[24]The possibility of renegotiation rules out some contracting features that might emerge otherwise, such as throwing away profit or not implementing profitable projects. The absence of renegotiation would also prevent parties from concealing information; for, the optimal contract would specify that no project is implemented unless two relevant projects are proposed and would costlessly ensure truthful revelation. This policy eliminates asymmetries of information. [To reintroduce real asymmetries of information in the absence of renegotiation, one may assume that each party observes with some probability the payoffs of only a subset of projects and therefore can claim that he is aware only of one relevant project.]

[25]That is, (y_3, z_3, e, E) solves:

$$\max\{B[E(1-e) + e(1-E)\alpha + eE(\alpha + (1-\alpha)y_3)] - g_P(E)\}$$
$$s.t.$$

$$E \in \arg\max\{B[E(1-e) + e(1-E)\alpha + eE(\alpha + (1-\alpha)y_3)] - g_P(E)\},$$

$$e \in \arg\max \quad \{b[E(1-e)\alpha + e(1-E) + eE(\alpha + (1-\alpha)z_3)] - g_A(e)\},$$
$$y_3 + z_3 = 1.$$

(although they are usually not considered in the literature), we conclude that the incomplete contracting approach cannot involve a loss of generality under these specific circumstances.

• What about third parties? Assuming that B is contractible (call it a profit), such parties (e.g., shareholders, headquarters) do exist in practice, who alter the principal's objective function by sharing her profit. The presence of third parties expands the contract space. A contract now specifies probabilities $\{p_k(m_P, m_A)\}$ of implementing projects as well as monetary rewards $\{B_P(m_P, m_A), B - B_P(m_P, m_A)\}$ to the principal and the third parties when profit is B. We assume that in case of renegotiation with the agent, the principal gets some fraction in $[0, 1]$ of the increase in profit brought about by the renegotiation.

We leave it to the reader to check that the *optimal complete contract in the presence of third parties* takes the following simple form: It specifies *a (random) authority relationship* (as in the absence of third parties) *cum a linear profit sharing scheme* in which the principal receives a fraction in $[0, 1]$ of the profit B. Profit sharing is therefore the new instrument, and serves to encourage initiative if the principal has too much incentives to monitor.[26] Because profit sharing schemes involving third parties are perfectly consistent with the incomplete contracting approach, we conclude that *the incomplete contracting approach cannot involve a loss of generality in the model of section 2.*

• We by no means want to argue that optimal complete contracts always boil down to simple authority relationships (possibly augmented by profit sharing schemes). In particular, if the agent responded to monetary incentives as in section 4.5, he could be rewarded simply for bringing information to the fore rather than by implementing the project he prefers, which is impossible in the standard incomplete contracting approach followed here. The ability to contract ex ante on projects then considerably alters the picture because it allows a disconnection between rewards for information acquisition and project choice. But we find it comforting that authority can under some circumstances be given the complete contracting interpretation of who gets his way in case of discordant announcements.

4 Factors favoring initiative when the principal has formal authority

4.1 Span of control, overload and initiative

It is often argued that the planning and allocation process of the large conglomerates that were formed in the '60s became bureaucratized, and that the headquarters were responsible

[26]The third parties' ex post expected profit is received ex ante by the principal through an auction of the shares. The reason why the third parties act as a "sink" rather than a "source" (to use Holmström (1982)'s terminology) is that from the principal's viewpoint, the principal's effort is never suboptimal when the principal receives the full profit B, but may be excessive in that it reduces the agent's initiative.

for too many units, whose strategy they could not understand or influence. This called for a refocus on "core businesses". The purpose of this subsection is to introduce the superior's span of control and overload into the analysis. Suppose that a superior has authority over m identical subordinates. Each subordinate i screens in a set of tasks as described in section 2 and learns the corresponding payoff structure with probability e_i. The principal's disutility of efforts is $g_P(\Sigma_i E_i)$ where E_i is the principal's probability of learning the payoffs of agent i's activity. The subordinates' tasks are independent. There is a fixed cost f per subordinate.[27] So, the principal's payoff is:

$$u_P = \Sigma_i[E_i B + (1 - E_i)e_i \alpha B - f] - g_P(\Sigma_i E_i). \tag{9}$$

Each agent's reaction curve is still given by

$$(1 - E_i)b = g'_A(e_i). \tag{10}$$

We assume that the equilibrium is symmetric[28] and stable.[29]

$$(1 - \alpha e)B = g'_P(mE), \tag{11}$$

and

$$(1 - E)b = g'_A(e). \tag{12}$$

Let $\{E(m), e(m)\}$ denote the solution of the system of equations $\{(11),(12)\}$. Abusing notation, let

$$u_P(m) \equiv mR(E(m), e(m)) - g_P(mE(m)),$$

where

$$R(E(m), e(m)) \equiv E(m)B + [1 - E(m)]e(m)\alpha B - f$$

is the revenue per subordinate. Using the envelope theorem and treating m as a real number, the optimal span of control is obtained from

$$\frac{du_P}{dm} = [R(E(m), e(m)) - E(m)g'_P(mE(m))] + m\frac{\partial R}{\partial e}\frac{de}{dm} = 0. \tag{13}$$

The expression in brackets in (13) is the marginal profit associated with a unit increase in the span of control. An extra agent brings revenue R, but requires attention E which raises

[27]The superior would choose to have an infinite number of subordinates in the absence of a fixed cost (or, equivalently, of a positive reservation wage of the subordinates.) A finite size is obtained when $f > [g'^{-1}_A(b)]\alpha B$.

[28]There also exist asymmetric equilibria in which the principal devotes all her attention to a subset of agents, who therefore lack initiative, and none to the others. To eliminate asymmetric equilibria, one can assume that the probabilities $\{E_i\}$ are sufficiently nonsubstitutable in the principal's disutility of effort function.

[29]In the quadratic case (see the example in subsection 3.3), this amounts to $\alpha b B < m$.

the cost of supervision by Eg'_P, the "overload cost". The second term, in $\frac{\partial R}{\partial e}\frac{de}{dm} > 0$, is the "initiative effect", and measures the increase in the agents' effort associated with a reduction in oversight.

We will say that a firm is in a situation of *overload* if the marginal profit of an extra employee, keeping employee behavior constant, is negative. Equation (13) shows that *it is always optimal for the firm to be in a situation of overload* so as to credibly commit to reward initiative.[30]

Remark 1: The analysis in this subsection has an interesting dynamic application: Suppose that the implementation of projects takes place continuously over time, and that at each point in time the principal can freely adjust the span of control by hiring or firing subordinates. Assume furthermore that the principal acquires experience about her subordinates as time passes by (there is learning-by-doing in monitoring). Then, the above trade-off between overload costs and initiative has the following dynamic equivalent: Letting the firm *grow fast* (that is, hiring new subordinates at a high speed) involves high overload costs and therefore a loss of control for the principal; on the other hand, a *slow growth* policy is more likely to stifle the subordinates' initiative as the principal acquires experience on monitoring them.

Remark 2: We have assumed that all agents are subordinates. More generally, the extent of "partial integration" determines the level of overload, with more agents subject to the principal's authority corresponding to more overload.

Remark 3: As noted in subsection 3.4, another credible way to promoting initiative is for the principal to share B, provided B is verifiable, with a third party. [Recall that division managers or CEOs are not residual claimants in practice.] While it has the drawback of reducing the principal's incentives on other tasks as well, profit sharing may reduce the desirability of overload.

Remark 4: Overload is one way of raising the marginal cost of monitoring. Another way consists in refraining from investing in the monitoring structure (on this, see also the next subsection.) A case in point is the relationship between Fujitsu and its British subsidiary ICL. Fujitsu wanted to commit to preserve ICL's independence and initiative after acquiring it in 1990 and seems to have been successful in this respect.[31] Of particular interest is the fact that only two Japanese managers are resident in London. [Fujitsu also wants to float

[30] Because the marginal profit is negative, the principal would be better off committing herself, say, to playing golf rather than reaching overload. The problem with this is that playing golf is not a credible commitment (recall that the $g_P(\cdot)$ function summarizes the principal's disutility of supervision and therefore already includes the cost of forgone opportunities). Overload is a credible commitment not to stifle initiative.

[31] See *The Economist* (Management Focus. April 10, 1993).

shares of ICL in 1995 while retaining a majority holding, in order to reduce its stake and further guarantee ICL's independence. On this, see remark 3 above.]

4.2 More lenient rules

The purpose of this subsection is to provide a link between the distribution of decision rights and the existence of rules. The concept of "rule" has been given several meanings in the literature. Here, we will define a rule as a *contractual constraint on the agent's action space.*[32] More concretely, we will assume that the set of projects that can be screened by the agent can ex ante be divided into two subsets N and $\tilde{N}$. The agent will screen projects either in N or in $\tilde{N}$, but not in both. There are therefore three contractual possibilities for a given allocation of authority: two rules ("the agent must screen in N", "the agent must screen in $\tilde{N}$"), and the lack of rule, under which the agent is free to pick his screening strategy. Naturally, one of the rules, namely the one that forces the agent to pick the subset he would choose by himself, is equivalent to the lack of rule. So, there are really two choices: constrain or not constrain the agent by a rule.

Concretely, we will assume that the two candidate research strategies N and $\tilde{N}$ differ in only one respect: The principal has less expertise on research strategy $\tilde{N}$ than on N and therefore has a higher cost of monitoring projects in $\tilde{N}$ than in N. We then ask, is a rule more likely to be imposed when the principal has formal authority?

Our main result is that *a rule is more likely to be imposed when the principal has formal authority, that is within an organization rather than across organizations*, for the following reason. An independent agent is not threatened by the principal's being well informed, because the principal cannot overrule him. Actually, in the simultaneous model studied here, the agent even strictly gains by choosing research strategy N, because the principal is more likely to come up with a useful idea. In contrast, in the absence of rule, a subordinate may well choose research strategy $\tilde{N}$ in order to protect himself against interference from his superior. Whether the superior should allow her subordinate to do so hinges on which of initiative and loss of control is the more important concern. If the superior has very little expertise in $\tilde{N}$ and congruence is low, the superior wants to force the agent to adopt research strategy N.

So, let us assume that the principal's marginal cost of investigating recommendations in $\tilde{N}$, $\tilde{g}'_P(E)$, is higher than that of investigating recommendations in N, $g'_P(E)$. Whether the agent searches in N or $\tilde{N}$ is contractible. In each research line (N or $\tilde{N}$), there are as earlier

[32] As we will predict, such contractual restrictions are mainly enforced *within* organizations, even though they are sometimes observed *across* organizations, for example in research contracts between independent laboratories and product manufacturers, which sometimes restrict the research line to be pursued by the laboratory.

two relevant projects and a bunch of negative value projects. The agent can investigate only one research line.

Consider first the case of *P-formal authority.* Having the subordinate screen within $\tilde{N}$ instead of N shifts the superior's reaction curve downwards without affecting the subordinate's reaction curve. The equilibrium efforts therefore satisfy:

$$E(\tilde{N}) < E(N) \quad \text{and} \quad e(\tilde{N}) > e(N).$$

As expected, the subordinate demonstrates more initiative and the principal has less control when the superior has a higher monitoring cost. One way for the principal to commit to a lower degree of intervention is thus to let the agent choose whether to screen within N or $\tilde{N}$. Furthermore, the agent, if given the freedom, chooses $\tilde{N}$ in order to limit the threat of intervention, at least if congruence is low.[33]

On the other hand, a lack of congruence makes initiative less relevant because the subordinate most often recommends projects that are useless to the superior. The superior therefore prefers to impose a rule when the congruence parameter is small ($u_P(N) > u_P(\tilde{N})$).[34]

Consider now the *case of A-formal authority.* An independent agent is not threatened by the principal's being well informed, because the principal cannot overrule him. Actually, in the simultaneous model studied here, the agent even strictly gains by choosing research strategy N, because the principal is more likely to come up with a useful idea.[35] Rules therefore are therefore never imposed upon an independent agent.[36]

---

[33]In the absence of a rule, the subordinate may not want to inflict high monitoring costs on the superior if congruence is high. It may turn out that he benefits from the superior's being well informed. But the subordinate definitely prefers research strategy $\tilde{N}$ if the congruence parameter α is low. To see this, let us index the principal's disutility-of-effort function by a parameter $K \in [0,1], h_P(E,K)$, with $h_P(E,0) = \tilde{g}_P(E), h_P(E,1) = g_P(E)$ and $h'_P(E,K)$ decreasing with K. Then, by the envelope theorem:

$$\frac{du_A}{dK} = (\alpha - e)b\frac{dE}{dK},$$

which is unambiguously negative for α close to zero since $\frac{dE}{dK} > 0$ and $\lim_{\alpha \to 0} e = g'^{-1}_A[(1 - g'^{-1}_P(B))b] > 0$.

[34]As in the previous footnote, let us index the principal's disutility-of-effort function by a parameter K, $h_P(E,K)$, with $h_P(E,0) = \tilde{g}_P(E), h_P(E,1) = g_P(E)$, and both $h_P(E,K)$ and $h'_P(E,K)$ decreasing with K ($h_P(0,K) = 0$ for all K and $h'_P(E,K)$ decreasing in K imply that $h_P(E,K)$ is also decreasing in K). The envelope theorem implies that:

$$\frac{du_P}{dK} = (1-E)\alpha B\frac{\partial e}{\partial K} - \frac{\partial h_P(E,K)}{\partial K} > 0 \quad \text{for } \alpha \text{ close to zero.}$$

Thus, for α close to zero:

$$u_P(N) = u_P(K=1) > u_P(K=0) = u_P(\tilde{N}).$$

[35]More formally, since $E^d(N) > E^d(\tilde{N})$ and since for any e the function $E \mapsto u^d_A(E,e)$ is increasing in E, we have $u^d_A(N) = \max_e u^d_A(E^d(N),e) > \max_e u^d_A(E^d(\tilde{N}),e) = u^d_A(\tilde{N})$.

[36]Rules could arise under A-formal authority for instance if N and $\tilde{N}$ differed not in the principal's ability to monitor them, but in the payoffs.

4.3 Urgency and delegation

It is sometimes observed that the need to adapt quickly to customer requirements has forced firms to decentralize decision-making.[37] The purpose of this subsection is two-fold: It offers some insights on how delegation might be affected by the urgency of the decision and it illustrates the sequential case. We will formalize the urgency of the decision by the length of a product life cycle, but several alternative interpretations are possible. Suppose that the superior can investigate only once the project proposal has been made. Let T denote the horizon, that is the time elapsed between the proposal (date 0) and the date at which the product becomes obsolete. Abusing terminology, we will let T stand for the product life cycle. A congruent project yields profit B per unit of time between the starting date for production $t \geq 0$ and date T, at which time a superior substitute arrives on the market. A noncongruent project generates no profit. Similarly, the subordinate's preferred project gives him private benefit b per unit of time, between t and T; the other relevant project yields no private benefit.

The principal's decision problem consists in choosing a stopping time $S \in [0,T]$ at which to start production even if her investigations have not been successful by then. Waiting longer gives her more time for monitoring; namely, the probability that the principal learns the payoffs herself before some date τ, $F(\tau)$, is increasing, with density $f(\tau)$. There are however decreasing returns in monitoring, so the hazard rate $f(\tau)/[1-F(\tau)]$ is decreasing. For the first time in this paper, it makes a difference whether information is hard or soft. The case of hard information is trivial and uninteresting. If the project is congruent, the principal learns it from the agent and implements it immediately ($S=0$); on the other hand, in case of noncongruence, the superior never implements the subordinate's preferred project, because there is always some hope that she will discover her preferred one ($S=T$).[38] So let us assume that information is *soft*. For a given stopping rule S, the principal obtains flow profit B between the date of learning t and T if she learns payoffs at $t<S$, and has expected flow profit αB between S and T if she has not learned payoffs by date S and thus rubber-stamps the agent's project at date S. The principal's utility, *conditional* on the agent's having proposed a project, is therefore:[39]

$$\hat{u}_P = B\int_0^S [\frac{e^{-rt}-e^{-rT}}{r}]f(t)dt + \alpha B[1-F(S)][\frac{e^{-rS}-e^{-rT}}{r}],$$

where r is the principal's rate of time preference. This objective function is quasi-concave,

[37] See, e.g., the discussions of Wyman-Gordon and WalMart in Continental Bank (1993).

[38] This would not be so if the subordinate's preferred project yielded a strictly positive profit to the superior.

[39] In this formulation, the principal's cost of investigating is simply forgone profit due to delayed introduction of the product. The sequential model can also be formulated with a more standard disutility of the principal's effort: See footnote 17.

and the optimal stopping time either is zero if $\frac{\alpha}{1-\alpha} \geq \frac{f(0)}{1-F(0)}\frac{1-e^{-rT}}{r}$ (for α large enough, the principal rubber-stamps without even checking), or is given by the first-order condition:

$$\alpha = \frac{f(S)}{1-F(S)}(1-\alpha)[\frac{1-e^{-r(T-S)}}{r}]. \tag{14}$$

The left-hand side of (14) is the marginal cost of delaying the introduction of the product (divided by B); its right-hand side is equal to the conditional density of discovering the payoffs times the value of overruling the agent's choice between S and T (divided by B). The optimal stopping time if strictly positive increases with T ($1 > \partial S/\partial T > 0$) and decreases with α ($\partial S/\partial \alpha < 0$), as we would expect. Our main result is that for a short horizon, the principal conducts a cursory investigation. That is, *the principal is more likely to rubber-stamp, the more urgent the decision.*[40]

Last, we have been silent about the agent's behavior in that we have implicitly assumed that his search time (which could be random) was exogenously given. In general this time could depend on the urgency of the decision due to altered incentives of the agent. Another interesting question (in a world of random time of acquisition of information by the agent) is whether the agent would ever want to delay a proposal. Delaying the proposal delays the date of adoption (recall that $\partial S/\partial T < 1$), but also reduces the probability of being overruled (because $\partial S/\partial T > 0$). Clearly, an agent with a congruent project would not want to delay the proposal, but an agent with a noncongruent project might. In this case, a late proposal could signal a noncongruent project and be given substantial attention by the principal (since $\partial S/\partial \alpha < 0$). We conjecture that this would not reverse our main insight of positive correlation between urgency and rubber-stamping,[41] but a formal treatment lies outside the scope of this paper.

4.4 Reputation and random delegation

As is usual, an alternative to contracting or authority allocation is reputation. In practice, superiors try to develop reputations for "not intervening too often", or in a context with a larger project diversity, for "intervening only when justified". For conciseness, we will not develop a formal model of reputation building, but it is straightforward to do so along the familiar lines. We can sketch the broader idea of "intervening only when justified".

[40] Bolton and Farrell (1990) find that urgent decisions are more likely to be centralized. However the notion of centralization and the set of issues studied there are quite different from the ones analyzed here. In our terminology, Bolton and Farrell's "centralization vs decentralization" refers more to the allocation of formal authority than to the degree of the subordinate's real authority. Bolton and Farrell analyze a multi-agent investment problem and contrast the duplication-and-delay inefficiency of decentralization with the incompetence of a bumbling central decision maker.

[41] No "type" of agent would choose to delay the proposal if this also increased the principal's length of investigation.

Suppose that the superior faces a sequence of agents. For each agent, the payoff structure is as described above except that with some probability a noncongruent project imposes a nonnegligible loss on the principal instead of yielding profit zero. For incentives purposes, it may then be optimal for the superior to commit to overrule the agent only if the noncongruent project yields a negative profit, in that overruling in the other case is ex post optimal for the principal but reduces initiative too much to be worth it (A-formal authority would be optimal in section 3.3). A patient superior facing enough subordinates may then develop a reputation for overruling agents only if the noncongruent project yields a negative profit. So, *the superior uses her authority to overrule the subordinate "in important matters", but voluntarily relinquishes this authority (which is different from rubber-stamping) in matters that are less important to her.* This behavior would not be credible in a one-shot situation, in which the superior would systematically overrule if informed.

4.5 Performance measurement and subordinate's responsiveness to monetary incentives

The economics literature has emphasized the effect of the allocation of control on incentives. This subsection shows that incentives feed back on control. To this purpose, we generalize our theory to allow the agent to respond to monetary incentives. The profit is verifiable and the agent's utility for project k is $u(w) + b_k$ (where $u(0) = 0$, $u' > 0$, $u'' < 0$). Without loss of generality, the agent receives $w \geq 0$ when the principal's profit is B, and 0 otherwise.

The principal's net profit in her preferred project is now $\tilde{B} \equiv B - w$. The agent's average gain from being informed and having real authority is $\tilde{b} \equiv b + \alpha u(w)$ for $u(w) < b$, and $\tilde{b} \equiv u(w) + \alpha b$ for $u(w) \geq b$: When $u(w) < b$, the agent always picks his preferred decision; when congruent, the agent also receives wage w. The case $u(w) \geq b$ can be labelled "aligned incentives". The agent's monetary incentives are powerful enough that he forgoes his private benefit and always recommends the principal's preferred decision. Note that it is never optimal for the principal to set a wage just below $u^{-1}(b)$, because she can obtain congruent decision-making by raising the wage slightly.

The first-order conditions (5) and (6) under *P-formal authority* become:

$$(1 - \alpha e)\tilde{B} = g'_P(E), \tag{15}$$

and

$$(1 - E)\tilde{b} = g'_A(e). \tag{16}$$

The main conclusion of this section can be drawn from these two equations. A *higher wage increases real authority for two reasons*: First, by raising the agent's incentives, it makes it more likely that the agent will be able to recommend a project. Second, it reduces

the principal's incentive to monitor and therefore the probability that the principal overrules the agent.

Letting $\{E(w), e(w)\}$ denote the solution of {(15), (16)}, the derivative of the principal's profit with respect to w is:

$$\frac{du_P}{dw} = (1-E)\alpha(B-w)\frac{de}{dw} - [E + (1-E)e\alpha]. \tag{17}$$

The first term on the right hand side of (17) corresponds to the increase in initiative. The second term reflects the increase in the wage bill. The optimal wage (when incentives are not aligned) is equal to zero if α is small, but can be positive in general.

Remark 1: Chandler (in Continental Bank, 1993, p54-58) argues that decentralization combined with financial controls is effective in mature industries while the corporate office must be closely involved in industries in which new product development is critical. The application of this idea within a firm yields the so-called *dual top management approach*, which consists in "applying decentralized financial controls to mature businesses and a more centralized form of "strategic" control to capital-intensive, high-tech businesses". He observes that at (generally deemed successful) GE during the '80s, the managers of the "core" businesses - the long-established, mature businesses - received little planning or attention from the corporate office, and were run instead through strict monetary incentives (budgets and budget-based bonuses). The corporate office in contrast was very involved in the high-tech businesses (aerospace, aircraft engines, medical equipment), for which monetary incentives are harder to design (due to the uncertainty and the novelty of the products.) Our argument that better performance measurement raises an agent's real authority offers a rationale for the dual top management approach.

Remark 2: The equilibrium described above may not be immune to the possibility of renegotiation. For, suppose that the principal and the agent have ex ante agreed on a wage w and that the agent has learned the payoffs while the principal has not. Suppose that $b > u(w)$. Suppose further that the agent's information is hard information.[42] In case of noncongruence, the principal must raise the wage to $w^* = u^{-1}(b)$ in order to get a profit. She will be willing to do so if $w^* < B$. While renegotiation occurs, the principal may not want to commit to wage w^* ex ante, because she can get away with a lower wage when she herself is informed or when the projects are congruent. The analysis is otherwise qualitatively similar to that developed in the absence of renegotiation.

[42]It is equally straightforward to study the case of soft information.

4.6 Multiple principals

Having multiple principals is generally believed to impact on an agent's behavior. While providing a full treatment of this topic lies out of the scope of this paper, a short discussion already yields a number of useful observations. There are two dimensions to the deconcentration of ownership: returns and authority.

a) Splitting returns. The benefit B, provided it is monetary, can be split among several principals. Consider for instance the case of n equal partners (or "co-owners"), each entitled to receive B/n in case of success of the project. We assume the same cost function for all principals. The set of principals as a whole is informed if any of them is. Because they all want to maximize profit, the allocation of authority among them is irrelevant. Each principal's and the agent's reaction curves are respectively given by

$$(1-E)^{n-1}(1-e\alpha)\frac{B}{n} = g'_P(E) \tag{5'}$$

and

$$(1-E)^n b = g'_A(e). \tag{6'}$$

Spreading monetary benefits among several principals has two effects on initiative. First, it generates free riding and therefore reduces monitoring. This effect dominates when the principal's cost function is not too convex, as is the case for instance for a quadratic cost.[43] In this case, an increase in the number of principals raises initiative and results in a loss of control. On the other hand, with a very convex cost function, the multiplication of monitors substantially improves the monitoring structure, which may reduce initiative.[44]

Remark 1: We have assumed that the principals have similar monitoring abilities. Multiplying the number of monitors may increase monitoring if the principals' talents are complementary.[45]

---

[43]To show this, rewrite (5') and (6') in terms of the probability $\mathcal{E}$ that the principals be informed: Let $E(\mathcal{E}, n)$ be defined by

$$(1-E)^n = 1-\mathcal{E}.$$

The first-order conditions are then:

$$(1-\mathcal{E})(1-e\alpha)B = n(1-E(\mathcal{E},n))g'_P(E(\mathcal{E},n)) \tag{5''}$$

and

$$(1-\mathcal{E})b = g'_A(e). \tag{6''}$$

In the $(\mathcal{E}, e)$ space, an increase in n shifts only the principals' reaction curve, through a change in the right-hand side of (5''). We leave it to the reader to check that for a quadratic g_P function, the right-hand side of (5'') increases with n.

[44]This point is most easily demonstrated with the following functions: $g_P(E) = 0$ for $E \leq E_0$, and $= \infty$ for $E > E_0$. Then the probability that the principals are informed, $\mathcal{E}$, is given by $1-\mathcal{E} = (1-E_0)^n$.

[45]Consider a project that yields a strictly positive profit only if both its marketing and manufacturing sides are satisfactory. The use of two experts in the two fields increases the efficiency of the monitoring process.

Remark 2: The insights of this subsection should also apply to multi-layers hierarchies where, for instance, downstream agents are monitored by both a principal and a supervisor (middle-man). The monetary benefits are then spread among several (upper) layers. A systematic analysis of complex organizations however lies beyond the scope of this paper.

b) Splitting authority: Splitting authority among several principals obviously has no consequence if the principals' objectives are aligned as in the previous example. But authority is often split among principals with imperfectly aligned objectives (marketing and manufacturing divisions in a matrix organization, multiple ministries, chambers in congress, partners in a joint venture, creditors in a bankruptcy process). Who has real authority then hinges on the matrix of congruence parameters among principals and agent, as well as on the governance mechanism (for example, each principal can have veto power, or there can be majority voting with or without the participation of the agent.) Depending on these considerations, the agent's initiative may be enhanced or reduced by the split of authority. A conflict of interest among principals may increase the probability of veto by one of them. It may also raise each principal's incentive to become informed and not to rely on the other principals' recommendations. On the other hand, for more collegial decision processes, the agent may be able to "play" his multiple principals against each other and thereby get his way.[46]

5 Authority and communication

This section extends the basic framework by introducing the possibility that the agent communicates some prior information he may privately hold about the projects. A natural question then is whether the allocation of formal authority affects the communication of (relevant) information by the agent.

We assume that at the beginning the agent can communicate information that reduces the principal's marginal cost of investigation from g'_P to $\tilde{g}'_P$ such that $g'_P(E) > \tilde{g}'_P(E)$ for all $E > 0$. [For example, the agent might privately know that the two relevant projects belong to a subset N_1 of N and decide whether to reveal N_1 to the principal.] The action of communicating information to the principal is noncontractible. The timing is as described in section 2 except that the agent first chooses whether to communicate his private information to the principal. Then the two parties choose noncooperatively how much effort (E and e) to invest in learning the projects' payoffs. Depending on the allocation of formal authority the equilibrium efforts are given by the first-order conditions (5) through (8) for the relevant marginal-disutility-of-effort function for the principal ($g'_P(\cdot)$ or $\tilde{g}'_P(\cdot)$).

[46]See Davis-Lawrence (1977) for a description of such behaviors.

Communication shifts the principal's reaction curve up (regardless of the allocation of formal authority) and has no effect on the agent's reaction curve. The principal's monitoring effort E thus increases in a stable equilibrium. The question of whether the agent wants to communicate information to the principal thus boils down to whether the agent gains from the principal's being better informed. Without loss of generality let us index the principal's marginal disutility function by a communication parameter $K \in [0,1]$, $h'_P(E,K)$ with $h'_P(E,0) = g'_P(E)$, $h'_P(E,1) = \tilde{g}'_P(E)$ for all E and $h'_P(E,K)$ decreasing in K. We just noted that the principal's equilibrium efforts $E(K)$ and $E^d(K)$ increase with K. Using the envelope theorem, the impact of communication on the agent's utility (see equations (2) and (4)) is given by:

$$\frac{du_A}{dK} = (\alpha - e)b\frac{dE}{dK} \tag{18}$$

and

$$\frac{du_A^d}{dK} = (1 - e^d)\alpha b\frac{dE^d}{dK}. \tag{19}$$

An independent agent always benefits from the principal's being better informed. In contrast, a subordinate wants to communicate information if his expected gain from the superior becoming informed, αb, exceeds the expected benefit from having real authority, eb, or $\alpha > e$. When congruence is low, $\alpha < e$,[47] so the agent is better off not communicating his information. *In case of low congruence there is more communication by an independent agent.*

In this framework, there is actually always at least as much communication under A-formal authority. This may no longer be the case if congruence is high and the agent incurs a direct (fixed) cost of communicating the information, as can be seen from equation (19). The agent no longer derives a benefit from the principal's being ex ante well informed if he himself is well informed (e^d close to 1). If the agent's private benefit is high enough so that e^d is indeed close to 1, the independent agent does not bother incurring the cost of communicating ex ante information. In contrast, it may be the case (this can be checked with quadratic payoffs) that, provided α is high enough,

$$\frac{du_A}{dK} > \frac{du_A^d}{dK} \simeq 0. \tag{20}$$

To summarize this discussion, the allocation of formal authority affects the agent's incentives to communicate prior information to the principal. The impact of the allocation of formal authority on communication depends upon the parameters of the model, in particular the degree of congruence between the principal's and the agent's objectives. More communication may take place under P-formal authority if these objectives are sufficiently

[47] One has $\lim_{\alpha \to 0} e = g_A'^{-1}[(1 - g_P'^{-1}(B))]b > 0$.

congruent, less communication will take place if they are too dissonant. We conclude this section with two remarks:

Remark 1: The discussion above has focused on the effects of the allocation of *formal* authority on communication. One can reasonably conjecture that the organizational factors that affect the distribution of *real* authority under P-formal authority (see section 4) will thereby also affect the amount of communication within the same organization. Whether the equilibrium amount of communication varies "comonotonically" with the amount of initiative left to the agent by the principal through various commitment devices of the kind mentioned in section 4 (span of control, multiple principals, etc.), is left for future research.

Remark 2: The above analysis shows that communication can *never* be detrimental to an independent agent. Such a strong conclusion, however, is unlikely to be robust to various extensions of our basic model, in particular to the introduction of the agent's responsiveness to monetary incentives. For example, one could imagine that an independent agent might prefer not to help the principal find out that the two parties' preferences are congruent in order to credibly blackmail the principal ex post and thereby obtain a higher monetary compensation.

6 Summary and extensions

Let us first summarize our main points. In an organization, the structure and the distribution of information, the amount of communication among members, and the existence of rules all are endogenous and depend on the allocation of decision rights. Having the right to decide raises one's incentives to become informed; consequently vertical integration endogenously improves the principal's information about the agent's activity. Relatedly, it reduces the agent's initiative. It may also (although it need not) jeopardize communication by making the agent concerned about being overruled. Similarly, rules are more likely to be imposed within an organization than across organizations because integration makes the agent fearful of being overruled and therefore less trustworthy.

We also identified factors that may increase a subordinate's real authority: large span of control, lenient rules, urgency, reputation for moderate interventionism, performance measurement and multiple principals. There doubtless are other factors, the investigation of which we leave for future research.

This paper aims only at being a first step toward a more general theory of authority and its delegation. There are many desirable extensions. For instance, we have been concerned with a single decision right. Organizations in general allocate multiple decision rights among their members. A fascinating question is whether there exist forces that lead to the clustering

of certain types of decision rights in a single hand.

Another crucial extension is to allow for multi-layered hierarchies, and to study the complex webs of authority relationships in organizations. To illustrate the kind of questions to be answered here, suppose two subordinates in an organization collect information about potential projects and make different recommendations to their superior (the principal). Whose advice will the principal follow? That is, which agent has (real) authority over the other agent? The obvious answer is that *authority is delegated to the agent whose (possibly endogenous) preferences are most congruent with the principal's.* This point is well illustrated by the following hierarchy: The agent is as described in the two-tier model. The private benefit attached to his preferred project is b_A and his wage in case of success is w_A. Monitoring (which consists in learning the payoffs with probability E) is not performed by the principal, but by a supervisor who enjoys no private benefit and receives wage w_S from the principal in case of success. In this model, the supervisor clearly should be given authority, because his preferences with respect to project choice (maximize the probability of success) perfectly coincide with the principal's, even though his preferences over his monitoring effort differ from the principal's. Note that, more generally, whether the supervisor gets his way is issue contingent. The principal decides among the parties who collect information, in favor of the one whose objective is less biased by private benefits or can be most easily aligned with the principal's through formal incentives.[48]

Another question directly related to authority and its delegation is that of liability in hierarchies, and relates to a negative externality exerted on a third party by the choice of project. The existing legal literature on "vicarious" liability rules essentially argues that in situations of two-sided moral hazard where the externalities exerted by a hierarchy depend upon both a principal's and an agent's efforts, responsibility for damages to a third party should be shared between the two contracting parties. Even in situations where the agent alone can directly affect the third party, making him entirely liable is suboptimal if the principal can monitor and control the agent's level of care (that is, has authority) and the agent only has a limited ability to pay for damages. Thus, if the principal has a direct power of intervention, she should be held responsible as well (see Shavell (1987, chapter 7) for a statement of the argument). While providing some support for this basic principle (which has

[48]For example, a principal may arbitrate in favor of the agent when he feels the supervisor is biased (due to a prejudice against the agent, collusion with a rival agent, supervisor's own private benefit, and so forth).

Along the same lines, one can wonder whether the principal necessarily gains from choosing a "protégé" as a supervisor, that is someone with high congruence with her. It might be the case that the principal be better off with a supervisor sharing the same culture (congruent) with the agent even if this entails some cover-ups. The "protégé" might stifle initiative.

Our focus on private benefits does not preclude the existence of other factors influencing the allocation of authority such as the relative competency of the two parties. Who gets his way may also depend on the indispensability/bargaining power of the various parties, or on the desire of the organization to keep key personnel in the long run (Rotemberg (1993,1994)).

been applied, e.g., to the new environmental liability rules ("CERCLA") enacted in the US in 1980), our analysis suggests at least two further considerations for its enforcement: (i) First, the degree of liability imposed upon a principal in an integrated firm should vary with her ability to exert *real* authority: In particular, directors, parent companies or creditors should be held more liable if they are more directly involved in supervising the agent's activities (see Strasser-Rodosevich (1993) for an extensive account of the courts' views in this respect). We thus provide support for liability rules that can be made sufficiently flexible in order to accommodate various organizational characteristics that affect *real* authority. (ii) Second, our analysis suggests that liability rules which allocate much responsibility to principals in integrated structures may have undesirable consequences "ex ante" when we allow for an endogenous choice of the authority structure. Such liability rules may indeed induce excessive divestiture in situations where coordination considerations would naturally favor the emergence of integrated structures.[49] There is thus a trade-off between responsabilizing the owners or managers in integrated firms and at the same time avoiding inefficiencies in the allocation of formal authority.

The analysis of these and of other exciting questions related to authority and its delegation must await future research.

[49]Such phenomena appear to have occurred in the US following the introduction of the new environmental liability rules (see Ringleb-Wiggins (1990)).

References

Arrow, K. [1975], "Vertical Integration and Communication", *Bell Journal of Economics*, 6:173–183.

Bolton, P. and J. Farrell [1990] "Decentralization, Duplication, and Delay", *Journal of Political Economy*, 98:803-826.

Bolton, P. and M. Dewatripont [1992] "The Firm as a Communication Network", mimeo, LSE, and Université Libre de Bruxelles, forthcoming, *Quarterly Journal of Economics.*

Burkart, M. , Gromb, D. and F. Panunzi [1994] "Large Shareholders, Monitoring and Fiduciary Duty", mimeo, LSE, and MIT.

Continental Bank [1993] "Executive Roundtable on Corporate Strategy in the '90s", *Journal of Applied Corporate Finance*, 6:42–60.

Crémer, J. [1992], "Arm's-length Relationships", mimeo, IDEI, Toulouse.

Davis, S. and P. Lawrence [1977], *Matrix*, Reading: Addison-Wesley.

Grossman, S. and O. Hart [1986], "The Costs and Benefits of Ownership: A Theory of Lateral and Vertical Integration", *Journal of Political Economy*, 94: 691–719.

Hart, O. and J. Moore [1990], "Property Rights and the Theory of the Firm", *Journal of Political Economy*, 98: 1119–1158.

Holmström, B. [1979], "Moral Hazard and Observability", *Bell Journal of Economics*, 10:74–91.

Holmström, B. [1982], "Moral Hazard in Teams", *Bell Journal of Economics*, 13:324–340.

Marschak, J. and R. Radner [1972] *Economic Theory of Teams*, New Haven: Yale University Press.

Mirrlees, J. [1975], "The Theory of Moral Hazard and Unobservable Behaviour, Part I", mimeo, Nuffield College, Oxford.

Radner, R. [1992] "Hierarchy: The Economics of Managing", *Journal of Economic Literature*, 1382–1415.

Riordan, M. [1990], "Vertical Integration and the Strategic Management of the Enterprise", in *The Firm as a Nexus of Treaties*, Aoki, Gustafsson and Williamson, eds, Sage Publications.

Ringleb, A., and S. Wiggins [1990] "Liability and Large-Scale, Long-Term Hazards", *Journal of Political Economy*, 98:574-595.

Rodriguez, D. and D. Vayanos [1993], "Decentralization and the Management of Competition", mimeo, MIT.

Rotemberg, J. [1993], "Power in Profit Maximizing Organizations", mimeo, MIT.

Rotemberg, J. [1994]. "Formal Authority vs Power in Profit Maximizing Organizations", mimeo, MIT.

Sah, R., and J. Stiglitz [1986] "The Architecture of Economic Systems: Hierarchies and Polyarchies", *American Economic Review*, 76:716–727.

Schmidt, K [1991], "The Costs and Benefits of Privatization", mimeo, Universität Bonn.

Shavell, S. [1979], "Risk Sharing and Incentives in the Principal and Agent Relationship", *Bell Journal of Economics*, 10:55–73.

Shavell, S. [1987], *Economic Analysis of Accident Law*, Cambridge: Harvard University Press.

Simon, H. [1951] "A Formal Theory of the Employment Relationship", *Econometrica*, 19:293–305.

Strasser, K., and D. Rodosevich [1993] "Seeing the Forest for the Trees in CERCLA Liability", *Yale Journal on Regulation*, 11:493–560.

Williamson, O. [1975], *Market and Hierarchies: Analysis and Antitrust Implications*, New York: Free Press.

Williamson, O. [1985], *The Economic Institutions of Capitalism*, New York: Free Press.

Weber, M. [1968], *Economy and Society*, New York: Bedminster Press.